

Om protein - nucleic search, using frame_plus_p2n model

Run on: May 18, 2003, 07:51:53 ; Search time 2838 Seconds
(without alignments)

5342.693 Million cell updates/sec

Title: US-09-857-581-66
Perfect score: 2396
Sequence: 1 MLELALGLVIALPFRHNP.....AHSIVCVPLARIQVSKILLS 521

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh

0=/cgm2_1/USPTO_spool/US09857581/rnatc_12052003_084115_11910/app_query.fasta.1.7

11
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09857581_@CGN_1_1724_@rnatc_12052003_084115_11910 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6 -FCAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2270	94.7	1563	AF195812	AF195812 Pisum sat
2	2269	94.7	1774	AF195798	AF195798 Glycine m
3	2268	94.7	1566	AF195806	AF195806 Vigna rad
4	2268	94.7	1566	AF195807	AF195807 Vigna rad
5	2268	94.7	1566	AF195808	AF195808 Vigna rad
6	2268	94.7	1566	AF195810	AF195810 Trifolium
7	2268	94.7	1566	AF195811	AF195811 Trifolium
8	2267	94.6	1566	AF195809	AF195809 Vigna rad
9	2267	94.6	1722	AF135484	AF135484 Glycine m
10	2251	93.9	1824	AF022462	AF022462 Glycine m
11	2251	93.9	1824	AF195799	AF195799 Glycine m
12	2201.5	91.9	1902	AF195819	AF195819 Glycine m
13	2189	91.4	1800	AF195818	AF195818 Glycine m
14	2183	91.1	1501	AF195801	AF195801 Medicago
15	2180	91.0	1501	AF195800	AF195800 Medicago
16	2179	90.9	1501	AF195815	AF195815 Trifolium

QY 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyAsp*****ValAla 120
 DB 301 TTCACGACCAAGTTCACAACTCTGTAAGACGCTCACTTACGACCACTCTGGGCGC 360
 QY 121 *****Pro***GlyProTyTrp***PheValArgGlyLeuIleMetAsnAspLeu 140
 DB 361 ATGCTTCATTGCGACCTTACTGAAAGTTCGTGAGAGCTCATATGACGACCTTCTC 420
 QY 141 AsnAlaThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgGly***Leu 160
 DB 421 AACGCAACACGCTCAACAGCTCAGGCTTTGAGAGCCCAACACATCCGAAAGTTCTC 480
 QY 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGlnGluLeu 180
 DB 481 AGGCTTATGGCCCAAGCGCAGAGGCCCAAGCCCTTGCATCGTCAAGAGAGCTTCTC 540
 QY 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyLysGluGluIleArgAsp 200
 DB 541 AATGACCAACAGCAACCATCTCATGATGATGCTCGGCGAGGCTGAGAGATCAGAGAC 600
 QY 201 IleAlaArgGluValLeuLysIle***GlyGluTySerLeuThrAspPheIle***Pro 220
 DB 601 ATCGCTCCGACAGTCTTACAGATCTTGGCGAATACAGCTCACTGACTTCACTGCGCT 660
 QY 221 LeuLys***LeuLysValGlyLysTyArgLysArgIleAspAspIleLeuAsnLysPhe 240
 DB 661 TTGAAGATCTCAAGGTTGAAAGTATGAAAGAGATTTGATGACATTTGAACAACTTC 720
 QY 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
 DB 721 GACCTGCTGTTAAAGGTCATCAAGAGCGCGTGAATGATGATGAGAGAGAAAGAAC 780
 QY 261 GlyGlu*****GluGlyGlu***SerGlyVal***LeuAspThrLeuGluPheAla 280
 DB 781 GGAGAGATTTGTAAGGCGACGCGGCTCTTCTCGACACTTCTGTAATTCGT 840
 QY 281 GluAspGluThr***GluIleLysIleThrLys*****IleLysGlyLeuValAsp 300
 DB 841 GAGGACGAGCAGTGAATCAAAATTACCAAGAGCAATCAAGGCGCTTGTGTCGAC 900
 QY 301 ***PheSerAlaGly***AspSerThrAla***ThrGluTrpAlaLeuAlaGluLeu 320
 DB 901 TTTTCTCTCAGGAGCAGATTCACAGCGGTGCGACAGAGTGGCACTTGGCAGAGCTC 960
 QY 321 IleAsnAsnPro***ValLeu*****AlaArgGluGlu***TyrSerValValGlyLys 340
 DB 961 ATCAACAATCCAGGCTGTTCAAAAGGCTGTGAGAGGCTCTACAGAGTGTGGGCAAA 1020
 QY 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyIleArgAlaIleValLys 360
 DB 1021 GATGAGCTCTGACGAAGTGCACATCAAACTTCTTACATTTAGGCGCATTTGTGAAG 1080
 QY 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys 380
 DB 1081 GAGCATTCGAGATGACCCACCACTCCAGTGTGTAAGAAAGAGTGCACAGAGAGTGT 1140

QY 381 ***IleAsnGly***Val***ProGluGlyValAlaLeu*****PheAsnValTrpGluVal 400
 DB 1141 GAGATTAAATGAGTATGTATGATCCAGAGGAGCATTTGTTCTTTCAATGTTGGCAAGTA 1200
 QY 401 Gly***Asp***LysTyTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu 420
 DB 1201 GGAAGAGCCCAAAATCTGGGACAGACATCAAGATTTCCGTCCGAGAGGTTCTTGAA 1260
 QY 421 Thr***AlaGluGlyGluAla*****LeuAspLeuArgGly***HisPheGlnLeu 440
 DB 1261 ACTGGCGCTGAAGGAGAGAGGCGCTTGTGATCTTGAAGGGCCAGCATTTCCAACTCTC 1320
 QY 441 ProPheGlySerGlyArg***MetCysProGlyVal***LeuAlaThrSerGly***Ala 460
 DB 1321 CCATTGGGTCTGGAGAGAGATGTGCGCTGTCGATTTGGCTTACTTCAAGAAATGGCA 1380
 QY 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGlyGln 480
 DB 1381 ACACCTTTCGATCTTATTCATATGCTTTGACCTGCAAGATGCTGGGCTTCAAGACAA 1440
 QY 481 IleLeuLysGly***AspAlaLysValSerMetGluGluArgGlyValLeuThrValPro 500
 DB 1441 ATTTGAAAGGTACAGATCCAAAGTTAGCATGAGAGAGAGCTGGCTTACCGTTCCA 1500
 QY 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520
 DB 1501 AGGCGACATGTCTCGTTGTGTTGCTTCCACTTGAAGATCGGCGTTCATTAACCTCTT 1560
 QY 521 Ser 521
 DB 1561 TCT 1563

RESULT 2
 AF195798
 LOCUS AF195798 1774 bp mRNA linear PLN 16-FEB-2000
 DEFINITION Glycine max isoflavone synthase 1 (Ifs1) mRNA, complete cds.
 ACCESSION AF195798
 VERSION AF195798.1 GI:6979519
 KEYWORDS
 ORGANISM
 Glycine max.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 1774)
 REFERENCE
 AUTHORS Jung, W., Yu, O., Lau, S. M., O'Keefe, D. P., Odell, J., Fader, G. and
 McConigle, B.
 TITLE Identification and expression of isoflavone synthase, the key
 enzyme for biosynthesis of isoflavones in legumes
 JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)
 MEDLINE 20124255
 PUBMED 10657130
 REFERENCE 2 (bases 1 to 1774)
 AUTHORS Jung, W., Yu, O., Odell, J., Fader, G. and McConigle, B.
 TITLE Direct Submission
 JOURNAL Submitted (18-OCT-1999) Nutrition and Health, DuPont, P. O. Box

Db 301 TTCACACAGGTTCCAAACCTCAGCATAGACCTTCACTATGATGAGTGGCC 360
 Qy 121 *****Pro**GlyProTyrTrp***PheValArgLysLeuIleMetAsnAspLeuLeu 140
 Db 361 ATGTTCCCTTCGACCTTACTGAGAGTGTGGAGAAAGCTCATGATCAAGACCTTCC 420
 Qy 141 AsnAlaThrThrValAsn**LeuArgProLeuArgThrGlnGlnIleArgLys***Leu 160
 Db 421 AACGACACCATGTAACAGATTGAGGCTTTGAGACCCACAGACATCGGACATTCCTT 480
 Qy 161 Arg**MetAlaGln**AlaGluAla**LysProLeuAsp***ThrGluGluLeuLeu 180
 Db 481 AGGTTATGGCCCAAGGGCGAGAGGACAGAAAGCCCTTACCTTACCCAGAGACTTCG 540
 Qy 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyLysLysLysLysLysLys 200
 Db 541 AATATGACACAGACCATCTCATGATGATGATGATGATGATGATGATGATGATGATG 600
 Qy 201 IleAlaArgLysValLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 220
 Db 601 ATGCTGCGCAGAGGCTTCTTACATCTTGGCGAATACAGCCCTCACTGACCTTCACTG 660
 Qy 221 LeuLys**LeuLysValGlyLysLysLysLysLysLysLysLysLysLysLysLys 240
 Db 661 TTGAAGCATCTCAAGGTTGAGAGATGAGAGAGATGAGAGATGAGAGATGAGAGATG 720
 Qy 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
 Db 721 GACCTGCTGCTTGAAGAGATCATCAAGACCCCTGATGATGATGATGATGATGATG 780
 Qy 261 GlyLys****GluGlyGlyLys***SerGlyVal**LeuAspThrLeuLeuGluPheAla 280
 Db 781 GAGAGGTTGTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTG 840
 Qy 281 GluAspGluThr**GluIleLysIleThrLys****IleLysGlyLeuValAsp 300
 Db 841 GAGATGAGACATGAGATCAAAATCAACAGACCATCAAGGCTTGTGTGAGAC 900
 Qy 301 ***PheSerAlaGly**AspSerThrAla****ThrGluTrpAlaLeuAlaGluLeu 320
 Db 901 TTTTCTCGACAGAACAGACTTCACAGCGGTGGACACAGAGTGGCATTTGGCAACTC 960
 Qy 321 IleAsnAsnPro**ValLeu****AlaArgGluLys**TyrSerValValGlyLys 340
 Db 961 ATCAACATCTTAAAGGTGTGAAAGGCTGTGAGAGGCTTACAGTGTGTGGAAAG 1020
 Qy 341 Asp**LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360
 Db 1021 GACAGACTGTGAGAGAGATGACACTCAAAACCTTCTTACATTAGAGCAATCGTGAAG 1080
 Qy 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys 380
 Db 1081 GAGCATTTCCGACATGACCCCGACCTCCAGGTGTCAAAAGAAATGACAGAAAGTGT 1140
 Qy 381 ***IleAsnGly**Val**ProGluGlyLysLeu****PheAsnValTrpGlnVal 400
 Db 1141 GAGATTATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1200

Qy 401 Gly***Asp***LysTrpTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu 420
 Db 1201 GGAAGAGACCCCAATATATGAGAGACATCGAGATTCCTCTGAGAGGTTCTTGAAG 1260
 Qy 421 Thr***AlaGluGlyGluAla****LeuAspLeuArgGly***HisPheGlnLeuLeu 440
 Db 1261 ACGGGGCTTAAGGGAG 1320
 Qy 441 ProPheGlySerGlyArg**MetCysProGlyVal**LeuAlaThrSerGly***Ala 460
 Db 1321 CCATTGGGTGTGAG 1380
 Qy 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGly 480
 Db 1381 ACATCTTCGACATCTTATACATGCTTGAATGCTGCAAGTGTGCTGCAAGAGCAG 1440
 Qy 481 IleLeuLysGly**AspAlaLysValSerMetGluLysArgLysLysLysLysLys 500
 Db 1441 ATATTGAAGGAGTGTGATCCCAAGTTAGATGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 Qy 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520
 Db 1501 AGGCACTATGCTTGTGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 1560
 Qy 521 Ser 521
 Db 1561 TCT 1563

RESULT 4
 AF195807 1566 bp mRNA linear PLN 16-FEB-2000
 LOCUS AF195807
 DEFINITION Vigna radiata isoflavone synthase 2 (lfss2) mRNA, complete cds.
 ACCESSION AF195807
 VERSION AF195807.1 GI:6979537
 KEYWORDS
 SOURCE Vigna radiata.
 ORGANISM Vigna radiata.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.

REFERENCE 1 (bases 1 to 1566)
 AUTHORS Jung W., Yu O., Lau S.M., O'Keefe D.P., Odell J., Fader G. and McGonigle B.
 TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes
 JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)
 MEDLINE 20124255
 PUBMED 10657130

REFERENCE 2 (bases 1 to 1566)
 AUTHORS Jung W., Yu O., Odell J., Fader G. and McGonigle B.
 TITLE Direct Submision
 JOURNAL Submitted (18-OCT-1999) Nutrition and Health, Dupont, P.O. Box 80402, Wilmington, DE 19880-0402, USA
 FEATURES location/Qualifiers
 source 1..1566

/organism="Vigna radiata"
/db_xref="taxon:157791"
1..1566
/gene="1fs2"
1..1566
/note="cytochrome P450"
/product="isoflavone synthase 2"
/protein_id="AAFP34528.1"
/db_xref="GI:6979538"
/translation="MULTIALGILVALFLHRLPTPAKSKALRLHPSPKRLPF
IGHILKDLHYALIDLSKGGPLSYSGSPVTVASTBELFKFLQTHEATSPN
TRPTSAIRRLTYDSSVAMVPFGPTKFRKLIMDLAATTNKLRLTQITKFL
RAMAQGAQKPLDTEELKMTNSTISMMGLGAEEIRDLAEVLKIFGYSITDPI
NPKLKLKVGKEKRIIDILNKFPVVERVKKREIVRRKGEVGEVGLDTL
LEFADETEKIKTDHILKGLVDFPAGTISVAIEMALDELINPKVLEKAREEV
YSVQKRLVDEVDTONLPYRAIVKEFRHAPLPVVRKKTCECEINGVYIPKAL
ILFNVQVGRDPKYMRRSEPRERFLETGAGARPLDRLQGHQILLPROGRNCP
GNLATSGMTPLASLIQCFDVLGPQGLIKGDARKVMEERAGLTVPRASLVCV
PLARIGVASIKLS"

BASE COUNT 414 a 393 c 387 g 372 t
ORIGIN

Alignment Scores:
Pred. No.: 1.03e-259 Length: 1566
Score: 2268.00 Matches: 454
Percent Similarity: 87.14% Conservative: 0
Best Local Similarity: 87.14% Mismatches: 67
Query Match: 94.66% Indels: 0
DB: 8 Gaps: 0

US-09-857-581-66 (1-521) x AFI95807 (1-1566)

OY 1 MetLeuLeuGluLeuAlaLeuGlyLeu***ValLeuAlaLeuPhe***HisLeuArgPro 20
DB 1 ATGTTCCTGCAACTGCACTTGCTTTATGCTTGGCTCTGCTCTGCACTTGCGCTCC 60
OY 21 ThrPro***Ala***SerLeuAlaLeuArgHisLeuProAsnProProSerPro***Pro 40
DB 61 ACACCCACTGCMAATCAAAAGCACTTCGCCACTCCCAAAACCCACCAAGCCCAAGCCCT 120
OY 41 ArgLeuProPheIleGlyHis***HisLeuLeuLysAspLysLeuLeuHisIleValAla*** 60
DB 121 CGTTCCTTCCTTCAATGACACTTCATCTTAAAGACAACTTCCTCCACTGCGCTCC 180
OY 61 IleAspLeuSerLysLysHisGlyProLeuPheSer*****PheGlySerMetProThr 80
DB 181 ATCGACTCTCCAAAAAACAATGCTCCCTTATCTCTCTACTTGGCTCCATGCAACC 240
OY 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln*****GluAlaThrSer 100
DB 241 GTTGTTCCTCCACACCAAGATGTTCAAGCTTCCTCCCAACCAAGGCGCAACTTC 300
OY 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla 120
DB 301 TTCACACACAGGTTCCAACTCAGCCATTAAGACCTCACTATGATAGCTAGTGGCC 360

OY 121 *****Pro***GlyProTyrTrp***PheValArgLysLeuIleMetAsnAspLeuLeu 140
DB 361 ATGTTCCCTTCGACCTTCACTGCAAGTTCGTGAGAAAGCTCACTCAAGCAAGCTTCTCC 420
OY 141 AsnAlaThrThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgLys***Leu 160
DB 421 AACGCCACCACTGTAAACAGTGGCTTTGAGAGCCCAACACAGATCCGCAAGTTCCTT 480
OY 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeuLeu 180
DB 481 AGGCGTATCGGCCCAAGCGCAGAGCAAGAGCCCTTCACTTCAAGCAAGAGCTTCTCG 540
OY 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGluLysIleArgAsp 200
DB 541 AATGACCAACACAGCACTTCATCATGATGCTCGGCAAGGCTGAGAGATCAAGAC 600
OY 201 IleAlaArgGluValLeuLysIle***GlyGlyTyrSerLeuThrAspPheIle***Pro 220
DB 601 ATCGCTCCGCAAGGTTCTTAAGATCTTGGGGAATCAAGCTTCACTGACTTCACTGCGCA 660
OY 221 LeuLys***LeuLysValGlyLysTyrGluLysArgGluLeuAspIleLeuAsnLysPhe 240
DB 661 TTGAAGCATCTCAAGGTTGGAAGATGAGAGAGATCAAGCATCTTGAACAGTTC 720
OY 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
DB 721 GACCTGTCGTGTAAAGATCATCAAGAGCCCTGATCTCGAGAGAGAAAGAAC 780
OY 261 GlyLysLys***GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
DB 781 GAGAGGTTGTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTT 840
OY 281 GluAspGluThr***GluIleLysIleThrLys*****IleLysGlyLeuValAlaAsp 300
DB 841 GAGATGAGACATGAGATGAGATCAAAATCAACAGACACATCAAGGCTTGTGTGAC 900
OY 301 ***PheSerAlaGly***AspSerThrAla***ThrGluThrAlaLeuAlaGluLeu 320
DB 901 TTTTCTTCGCGAGAACAGACTCCACAGCGGTGGCAGACAGATGGGCAATGGCAATC 960
OY 321 IleAsnAsnPro***ValLeu***AlaArgGluGlu***TyrSerValValGlyLys 340
DB 961 ATCAACATCTCAAGGTTGTGAAAAGGCTCGTGAAGAGGTTCAAGTGTGTGGCAAG 1020
OY 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360
DB 1021 GACAGACTTGGACCAAGTTGACACTCAAAACCTTCCTTACATTAGACAAATGTGAAG 1080
OY 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys 380
DB 1081 GAGACATTCGCAATGACCCGCCACTCCAGTGTCAAAAGAAAGTGAACGAGAGGTGT 1140
OY 381 ***IleAsnGly***Val***ProGluGlyAlaLeu*****PheAsnValTrpGlnVal 400
DB 1141 GAGATTATGATATATATATATCCAGAGAGAGATTTGATCTTCAATGTATGCAAGTA 1200

QY 121 *****Pro***GlyProTyrTrp***PheValArgIleLeuIleMetAsnAspLeuLeu 140
 DB 361 ATGCTCCCTTGGACCTTACTGAAAGTTCTGTGAGAGAGCTCATATGACAGCTTCTTC 420
 QY 141 AsnAlaThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgLys***Leu 160
 DB 421 AACGACACCTGTAAACAGATTGAGCCTTGAAGACCCAAACATCCGAAAGTTCTT 480
 QY 161 Arg***MetAlaGln***AlaGlnAla***LysProLeuAsp***ThrGlnGlnLeu 180
 DB 481 AGGCTTATGGCCCAAGCGCAGAGGACAGAGCCCTTGAATGACAGAGAGCTTCTG 540
 QY 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyGlnAlaGlnGlnIleArgAsp 200
 DB 541 AAATGACCAACAGCAGCATCTGATGATGCTCGGCGAGGCTGAGAGAGCAAGAC 600
 QY 201 IleAlaArgIleValLeuLysIle***GlyGlnTyrSerLeuThrAspPheIle***Pro 220
 DB 601 ATGCTGCGAGGTTCTTAAAGATCTTGGCGAATACAGCTCTCATCTTCACTTGGCCA 660
 QY 221 LeuLys***LeuLysValGlyLysTyrGlyLysArgIleAspAspIleLeuAsnLysPhe 240
 DB 661 TTGAAGCATCTCAAGTTGAAAGTATGAGAGAGAGATGACAGATCTTGAACAAAGTTT 720
 QY 241 AspProValValGlnArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
 DB 721 GACCTGTCGTTGAAAGATCATCAAGAGCGCGTGAATGCTGAGAGAGAGAAAGAAC 780
 QY 261 GlyGlu*****GlyGlyGlu***SerGlyVal***LeuAspThrLeuGlnGlnPheAla 280
 DB 781 GAGAGGTTGTGAGGAGGTGACGGGGGTTTCTTGTGACATCTTCTGTAATTCGT 840
 QY 281 GluAspGluThr***GluIleLysIleThrLys*****IleLysGlyLeuValValAsp 300
 DB 841 GAGATGAGACACGAGATCAATCAATCAACAGACATCAAGAGGCTTGTGTGAC 900
 QY 301 ***PheSerAlaGly***AspSerThrAla*****ThrGluTrpAlaLeuAlaGlnLeu 320
 DB 901 TTTTCTCGGAGAGAAACAGACTCCAGCGGCGGCAACAGAGTGGCATTTGGCAGAACTC 960
 QY 321 IleAsnAsnPro***ValLeu*****AlaArgGlnGlu***TyrSerValValGlyLys 340
 DB 961 ATCAACATCTTAAGTGTGTGAAAGGCTGTGAGAGAGGCTTACAGATGTTGGGAAAG 1020
 QY 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360
 DB 1021 GACGAGCTTGTGAGAGAGTGTGACCTCAAAACCTTCTTACATTTAGCAGCAATCGGAAG 1080
 QY 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GlyGlnCys 380
 DB 1081 GAGCATTTCCGATGACCCGCGCATCCCGAGTGTCAAAAGAAAGTGTGCAAGAGAGTGT 1140
 QY 381 ***IleAsnGly***Val***ProGlyGlyAlaLeu*****PheAsnValTrpGlnVal 400
 DB 1141 GAGATTAATGAGATGATGATCCAGAGAGAGCATTTGATCTCTCAATGATATGGCAAGA 1200
 QY 401 Gly***Asp***LysTyrTrpAspArgProSerGlu***ArgProGluArgPheLeuGln 420

DB 1201 GGAAGAGACCCCAATATCTGGAGACAGACATCGAGATTCCGTCTTGAGAGGTTCTTGAAG 1260
 QY 421 Thr***AlaGlnGlyGlnAla*****LeuAspLeuArgGly***HisPheGlnLeuLeu 440
 DB 1261 ACAGGGGCTGAGAGGAG 1320
 QY 441 ProPheGlySerGlyArg***MetCysProGlyVal***LeuAlaThrSerGly***Ala 460
 DB 1321 CCATTTGGGTCTGCGAGAGAGAGATGTGCTCGAGAGATGTGCTACTTCCGGAATGGCA 1380
 QY 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGlyGln 480
 DB 1381 AACATCTTGATCATCTTATTCAGTGTGATTCAGATGCTGAGTGTGCAAGAGAGAGAG 1440
 QY 481 IleLeuLysGly***AspAlaLysValSerMetGlnGlnArgAlaGlyLeuThrValPro 500
 DB 1441 ATTTGAAAGGTTGTGAG 1500
 QY 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520
 DB 1501 AGGGCAGATAGTCTGTCTGTGTTCACCTTGCAAGATCGGCGTTCATTAATCTCTT 1560
 QY 521 Ser 521
 DB 1561 TCT 1563

RESULT 6
 AF195810
 LOCUS AF195810 1566 bp mRNA linear PLN 16-FEB-2000
 DEFINITION Trifolium pratense isoflavone synthase 1 (Ifs1) mRNA, complete cds.
 ACCESSION AF195810
 VERSION AF195810.1 GI:6979543
 KEYWORDS
 SOURCE Trifolium pratense.
 ORGANISM Trifolium pratense
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Trifolium.
 REFERENCE 1 (bases 1 to 1566)
 Jung, W., Yu, O., Lau, S.M., O'Keefe, D.P., Odell, J., Fader, G. and
 McConigle, B.
 Identification and expression of isoflavone synthase, the key
 enzyme for biosynthesis of isoflavones in legumes
 JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)
 MEDLINE 20124255
 PUBMED 10657130
 REFERENCE 2 (bases 1 to 1566)
 Fader, G., Yu, O., Odell, J., Fader, G. and McConigle, B.
 Direct Substitution
 Submitted (18-OCT-1999) Nutrition and Health, Dupont, P.O. Box
 80402, Wilmington, DE 19880-0402, USA
 FEATURES
 1. 1566
 /organism="Trifolium pratense"
 /db_xref="taxon:57577"

gene 1. .1566
/gene="1fsl"
CDS 1. .1566
/note="cytochrome P450"
/codon_start=1
/product="isoflavone synthase 1"
/protein_id="AAFI4531.1"
/db_xref="GI:6979544"

/translation="MLLELALGLLVLALFLRLPTTPAKSKALRLPNPSPKRLPLF
IGHHLAKDKLHYALIDLSKRGPLFLYFGSMPTVASTPELFLKLTQHEXTSFN
TRPOTSAIRRLTYDSSVAMVPDIPYMKFVRKLIIMDDLIIMTLNRLRLTQRIKFL
RVMAGCAEAKPLDLTEELIKTNSTISMMKIGAEIRIDIAEVLKXIFGYSITDPI
NPLKRLKVKYKTKRIDIDILNKFPDVVVRVYIKRKRIYRRKRGSEVSGVFLDTL
LSEFADETTIKITKDIKGLVDFPSAGTSTAVALMALNLINPKVLEKAREEV
YSGVKRLVDEVDTONLYIRALIVKTEFRMHPPLPVYKKECEBCEINGYVIEGAL
ILFNWQYGDPRKMDRSEPRPERPLETGAEGEARPLDRLQGHFOLLFPFGSRMPC
GNLATSGMATLIALSLIQCPDLQVIGPQGLIKGDKAKSMERBAGLVPRANSLVCP
PLAIRGVASLILSL"

BASE COUNT 416 a 393 c 387 g 370 t
ORIGIN

Alignment Scores:
Pred. No.: 1.03e-259 Length: 1566
Score: 2268.00 Matches: 454
Percent Similarity: 87.14% Conservative: 0
Best Local Similarity: 87.14% Mismatches: 67
Query Match: 94.66% Indels: 0
DB: 8 Gaps: 0

US-09-857-581-66 (1-521) x AF195810 (1-1566)

QY 1 MetLeuLeuGluLeuAlaLeuGlyLeu***ValLeuAlaLeuPhe***HisLeuArgPro 20
DB 1 ATGTTGCTGCACTTCGACTTGCTTTATGCTTTTGCTCTGTTCTGCACTTCGCTCCC 60
QY 21 ThrPro***Ala***SerLysAlaLeuArgHisLeuProAsnProProSerPro***Pro 40
DB 61 ACACCCACTGCAGAAATCAAAAGACCTTCGCAATCCCAACCAACCAAGCCCAAGCCT 120
QY 41 ArgLeuProPheHleGlyHis***HisLeuLeuLysAspLysLeuLeuHisTyrAla*** 60
DB 121 CGCTCTCCCTTCATAGACACCTTCATCTCTTAAAGACAACTCTCCACTACGCACTC 180
QY 61 IleAspLeuSerLysHisGlyProLeuPheSer*****PheGlySerMetProThr 80
DB 181 ATGACCTCTCCAAACATGATGCTCTTATCTCTCTACTTGGCTCCAGCCCAAC 240
QY 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln*****GluAlaThrSer 100
DB 241 GTTGCTGCTCCACACAGATTTCTCAAGCTCTTCTCCAAACGACAGAGGCACTTCC 300
QY 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla 120
DB 301 TTCAACACAGGTTCCAAACCTCAGCCATAGAAGCCCTCAGTATAGTCAAGTGGCC 360
QY 121 *****Pro***GlyProTyrTrp***PheValArgLysLeuLleMetAsnAspLeuLeu 140

DB 361 ATGTTCCCACTCGACCTTACTGAGAACTTGATGAGAAAGCTCATCTAGAACCTTCTC 420
QY 141 AsnAlaThrThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgLys***Leu 160
DB 421 AACGCCACCACTGTAACAAAGTTGAGGCTTTGAGAGCCCAACAGATCCGCAACTTCTT 480
QY 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeu 180
DB 481 AGGGTTAATGGCCCAAGCCGACAGAGCCAGAAAGCCCTTGACTTGGACGAGACTTCTG 540
QY 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGlnIleArgAsp 200
DB 541 AAATGGACCAACAGACCACTTCATGATGATGCTGGCGAGCTTGAAGATTAAGAGAC 600
QY 201 IleAlaArgGluValLeuLysIle***GlyGluTyrSerLeuThrAspPheIle***Pro 220
DB 601 ATGGCTCCGAGGTTCTTAAGTCTTTGGCGAATACAGCTCTCACTCACTTCACTTGGCCA 660
QY 221 LeuLys***LeuLysValGlyLysTyrGluLysArgIleAspAspIleLeuAsnLysPhe 240
DB 661 TTGAAGCATCTCAAGGTTGAAAGATGAGAGAGATCGACAGCATCTTGAACAAGTTC 720
QY 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
DB 721 GACCCCTGCTGTTGAAAGACTCATCAAGAGCCCTGAGATCTCGAGAGCAAGAAAGAC 780
QY 261 GlyGluLys***GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
DB 781 GGAGAGGTTGATAGAGGTGAGGTGAGCGGCTTCTTGAACCTTGGTGAATTCGCT 840
QY 281 GluAspGluThr***GluIleLysIleThrLys*****IleLysGlyLeuValValAsp 300
DB 841 GAGATGAGACACAGAGATCAAAATCACACAGACCATCAAGGCTCTTGTTCGAC 900
QY 301 ***PheSerAlaGly***AspSerThrAla***ThrGluTrpAlaLeuAlaGluLeu 320
DB 901 TTTTCTCGAGAGGACAGACTCCACAGCGGTGGCAGACAGTGGCATTTGGCAAGCTC 960
QY 321 IleAsnAsnPro***ValLeu*****AlaArgGluGlu***TyrSerValValGlyLys 340
DB 961 ATCAACAAATCCTAAGCTGTGGAAAGAGCTCGTGAAGAGTCAAGTGTGGGAAG 1020
QY 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360
DB 1021 GACAGACTTGGAGACAGAGTGAACCTCAAAACCTTCTTCACTTGAAGCATGTGAAG 1080
QY 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys 380
DB 1081 GAGACATTCGCGCATGACACCCGCCACTCCAGTGTCCAAAGAAAGATGACAGAAAGCTGT 1140
QY 381 ***IleAsnGly***Val***ProGluGlyAlaLeu*****PheAsnValTrpGlnVal 400
DB 1141 GAGATTAATGATATGATGATCCAGAGGAGCATTAATGCTTCAATGATGAGCAAGTA 1200
QY 401 Gly***Asp***LysTyrTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu 420

Db 361 ATGTTCCCTTCGGAAGCTTACTGGAAGTTCTGAGGAAGGCTCATCATGAGCAAGCTTCTC 420

Qy 141 AenAlaThrThyValAsn***LeuArgProLeuArgThiGlnIleAlaGly***Leu 160

Db 421 AAGCGACCACTGTAAACAGTTGAGGCTTTGAGAGCCCAACAGATCCGGAAGTTCTC 480

Qy 161 Arg***MetAlaGly***AlaGluAla***LysProLeuAsp***ThiGluGluLeuLeu 180

Db 481 AGGTTTATGACCCAGGCGAGGAGCAAGAAAGCCCTTACTGACCAAGAGTTCTTG 540

Qy 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyValAlaGluIleArgAsp 200

Db 541 AATGACCAACAGCAACATCTCATGATGATGCTCCGCGAGGCTGAGAGATCAAGAC 600

Qy 201 IleAlaArgIleValLeuLysIle***GlyLysIleSerLeuThrAspPheIle***Pro 220

Db 601 ATGCGCGCAGGTTCTTAAGATCTTTGGCAATACAGCTCACTGATCATCTGCGCA 660

Qy 221 LeuLys***LeuLysValGlyLysIleValLysArgIleAspAspIleLeuAsnLysPhe 240

Db 661 TTGAAGCATCTCAAGGTTGGAAGTATGAAAGAGATCCAGCATCTTGAACAAGTTTC 720

Qy 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260

Db 721 GAACCTGTGTTGAAGAGATCATCAAGAGCCGCTGATGCTGAGAGAGAAAGAAC 780

Qy 261 GlyLysLys***GluGlyGlu***SerGlyVal***LeuAspThrLeuGluPheAla 280

Db 781 GAGAGGTTGTGAGGAGGTGAGTCAAGCGGCTTTCTTACACTTCTGTTAGTTCCGT 840

Qy 281 GluAspGluThr***GluIleLysIleThrLys***IleLysGlyLeuValValAsp 300

Db 841 GAGATGAGACAGCAGAGATCAAAATCACAGACATCAAGAGGCTTGTGTGAGC 900

Qy 301 ***PheSerAlaGly***AspSerThrAla***ThrGluTrpAlaLeuAlaGluLeu 320

Db 901 TTTTCTCGCAGCAACAGACTCCACAGCGGTGCGCAACAGATGCGCATTTGCGCAACTC 960

Qy 321 IleAsnAsnPro***ValLeu***AlaArgGluGlu***TyrSerValValGlyLys 340

Db 961 ATCAACATCTTAAGGTGTGAAAGGCTCTGAGAGGTTCTACAGTGTGTGGAAG 1020

Qy 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360

Db 1021 GACAGACTGTGAGCAAGATTGACACTCAAACTTCTTACATTAGAGCATGTGGAAG 1080

Qy 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys 380

Db 1081 GAGACATTCGCGATGACCGCCCACTCCAGGTGTCAMAAAGAGTGACAGAAAGTGT 1140

Qy 381 ***IleAsnGly***Val***ProGluGlyAlaLeu***PheAsnValTrpGlnVal 400

Db 1141 GAGATTATGATGTGATCTCCAGAGGAGCATTTCTCTTCATGATGTGCAAGTA 1200

Qy 401 Gly***Asp***LysTyrTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu 420

Db 1201 GGAAGAGCCCAATATCTGGAACAGACATGCAAGTTCTCTGAGAGGTTCTTAAG 1260

Qy 421 Thr***AlaGluGlyGluAla***LeuAspLeuArgGly***HisPheGlnLeuLeu 440

Db 1261 AAGGAGGCTGAAGGGAGAACAGGCTTGTGATCTTACGGAGCAACATTTTCAACTTCTC 1320

Qy 441 ProPheGlySerGlyVal***MetCysProGlyVal***LeuAlaThrSerGly***Ala 460

Db 1321 CCAATTGGGTTCTGGAGAGAGATGTGCTTGAGAGCATCTGACTTCTCGGAAATGCA 1380

Qy 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValIleuGlyProGluGlyGln 480

Db 1381 AACCTTCTGATCTCTTATTCAGTGTGACTTTCAGATGCTGGGTCACAGAGACG 1440

Qy 481 IleLeuLysGly***AspAlaLysValSerMetGluLysArgAlaGlyLeuThrValPro 500

Db 1441 ATATTGAAGGTGTGTAACCCCAAGTTAGCATGGAAGAGGCGCGCTCATCTGTCCA 1500

Qy 501 ArgAlaHisSerLeuValCysValProLeuAlaAlaGlyIleGlyValAlaSerLysLeuLeu 520

Db 1501 AGGCACTACTGTCTTGTGTGTTCACCTTGCAAGATCGGCGCTTGCAATCTTCTT 1560

Qy 521 Ser 521

Db 1561 TCT 1563

RESULT 8

AF195809 1566 bp mRNA linear PLN 16-FEB-2000

LOCUS AF195809

DEFINITION Vigna radiata isoflavone synthase 4 (lf54) mRNA, complete cds.

ACCESSION AF195809

VERSION AF195809.1 GI:6979541

KEYWORDS

SOURCE

ORGANISM

Vigna radiata.

Vigna radiata

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.

REFERENCE 1 (bases 1 to 1566)

AUTHORS Jung,W., Yu,O., Lau,S.M., O'Keefe,D.P., Odell,J., Fader,G. and McConigle,B.

TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes

JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)

MEDLINE 20124255

PUBMED 10657130

REFERENCE 2 (bases 1 to 1566)

AUTHORS Jung,W., Yu,O., Odell,J., Fader,G. and McConigle,B.

TITLE Direct Submission

JOURNAL Submitted (18-OCT-1999) Nutrition and Health, Dufont, P.O. Box 80402, Wilmington, DE 19880-0402, USA

FEATURES

source

1..1566

/organism="Vigna radiata"

/db_xref="taxon:157791"

1..1566

/gene="lf54"

CDS

1..1566
/gene="ife4"
/note="cytochrome P450"
/codon_start=1
/product="isoflavone synthase 4"
/protein_id="AAF34530.1"
/db_xref="GI:6979542"
/translation="MLLEALGCLLVIALFLHRLPTTPAKSLRLPNPSPKRLP
ICHLHLKDLKLVIALIDLSKKHGPLSLYRQSMPTVASTELFKLPLQTHATSFN
TRFQTSARILTYDSSVAMVPGPYRKINMDILNATVNLRLPLRTOQIRKFL
RVMAQGAQKPLDLTEELKMTNSTISMMMLGEAEIRIDIAEVYAKIFGEYSLDPI
WPLKHLVGVKERRIDILINKPDPVVERIVKRRRIYRRKNGEVEGSGVLDLTL
LEFADETMEIKITTDHFKGLVDFPSAGTSTAEATWALMELINPKVLEKAEV
YSVVGKORLVDEVDQNLPIYRAIVKETRMHPPPLPVKKKCTECSEINGVYIPEGAL
ILFNWQVGDPIKYMDRSEPERERLETGAEGRALDRLQGHFOLLPGSGRRMCP
GVNLATSGMATTILASLIQCFDLQVLPQQLIKGDAAKYSMERAGLTVRAHSIVCV
PLARIGVASKLS"

BASE COUNT 417 a 392 c 385 g 372 t
ORIGIN

Alignment Scores:

Pred. No.:	1,356-259	Length:	1566
Score:	2267.00	Matches:	454
Percent Similarity:	87.14%	Conservative:	0
Best Local Similarity:	87.14%	Mismatches:	67
Query Match:	94.62%	Indels:	0
DB:	8	Gaps:	0

US-09-857-581-66 (1-521) x AF195809 (1-1566)

QY 1 MetLeuLeuGluLeuAlaLeuGlyLeu***ValLeuAlaLeuPhe***HisLeuArgPro 20
DB 1 ATGTTCGCTGGAACCTTGACCTTGTTATGCTTTGGCTCTGTTCTGCACTTGCTGCC 60
QY 21 ThrPro***Ala***SerLyseAlaLeuArgHisLeuProAsnProProSerPro***Pro 40
DB 61 ACACCCCACTGCMAAAATCAAAAGCACTCCCATCTCCCAACCCACGACCAAGCCAAAGCCT 120
QY 41 ArgLeuProPheIleGlyHis***HisLeuLeuLysAspLysLeuLeuHisTyrAla*** 60
DB 121 CGTCTTCCTTCCTTCATAGACACCTTCATCTTAAAGACAACTTCTCACTAGCACTGC 180
QY 61 IleAspLeuSerLysLysHisGlyProLeuPheSer****PheGlySerMetProThr 80
DB 181 ATGACCTCTCCAAAAACATGCTTCCTTATTTCTCTCTACTTTGGCTCCATGCCAACCC 240
QY 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln****GluAlaThrSer 100
DB 241 GTTGTGCTCCACACACAGATTTGTTCAAGCTCTTCTCCAAACGACAGAGGCACTTCC 300
QY 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla 120
DB 301 TTCACACCAAGGTTCCAAACCTCAGCCATAGACGCTCACTATGATAGCTCAGTGCC 360
QY 121 *****Pro***GlyProTyrTrp***PheValArgLysLeuLeuMetAsnAspLeu 140
DB 361 ATGTTCCTTCGACCTTACTGGAAGTTCTGTGAGGAAGCTCATCATGACGACCTTCTC 420

QY 141 AsnAlaThrThrValAsn***LeuArgProLeuArgHisGlnGlnIleArgLys***Leu 160
DB 421 AAGGACCACTGTAAACAGAGTGAAGCCCTTGAGGACCCACAGATCCGACAGTCTTCT 480
QY 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeuLeu 180
DB 481 AGGATTATGCGCCACAGCGACAGGACACAGAGCCCTTCACTTGAACGAGACCTTCTG 540
QY 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGluLysIleArgAsp 200
DB 541 AATATGACCAACAGACACCACTCCATATGATATGCTCCGACGAGCTGAGAGATACAGAC 600
QY 201 IleAlaArgGluValIleuLysIle***GlyGluTyrSerLeuThrAspPheIle***Pro 220
DB 601 ATGCTCCGACGAGTCTTAAATGATCTTGGGGAATACAGCCCTCACTGACTCATGCGCA 660
QY 221 LeuLys***LeuLysValGlyLysTyrGlyLysArgGluIleAspAspIleLeuAsnLysPhe 240
DB 661 TTGAAGCATCTCAAGCGTTGGAAGCTATGAGAGAGATGACGACGATCTTGAACAGTTTC 720
QY 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
DB 721 GACCTGCTGCTTAAGAGGTATCAAGAGCCCTGTAGATCTGAGAGAGAGAAAGAAC 780
QY 261 GlyGlu*****GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
DB 781 GGAAGAGCTTGTGAGGTAGGTAGGTAGGTGTTTCTTCTTACACTTGTGTAATTCGCT 840
QY 281 GluAspGluThr***GluLysLysIleThrLys*****IleLysGlyLeuValValAsp 300
DB 841 GAGATAGAGACATGAGATCAAAATCAACAGACCAATCAAGAGGCTTGTGTGAC 900
QY 301 ***PheSerAlaGly***AspSerThrAla*****ThrGluTrpAlaLeuAlaGluLeu 320
DB 901 TTTTCTCGGAGGAAACAGACTCCACAGCGGAGGCAAGAGTGGCAATGGCAATCTC 960
QY 321 IleAsnAsnPro***ValLeu*****AlaArgGluGlu***TyrSerValValGlyLys 340
DB 961 ATCAACAAATCTTAAGGTGTGGAAGAAAGCTCGTAGAGGCTTACAGTGTGTGGAAAG 1020
QY 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360
DB 1021 GACAGACTGTGACAGCAAGTTCACACTCAAAACCTTCTTACATTAGACATGTGAAAG 1080
QY 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluGluCys 380
DB 1081 GAGACATTCGCAATGACACCGGCACTCCGAGTGTCAAAAGAAAGTGCAGAAAGTGT 1140
QY 381 ***IleAsnGly***Val***ProGluGlyAlaLeu*****PheAsnValTrpGluVal 400
DB 1141 GAGATTAAATGATATGATGATCCAGAGGAGCATTTGCTTTCATATGATATGCAAGTA 1200
QY 401 Gly***Asp***LysTyrTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu 420
DB 1201 GGAAGAGACCCCAATATCTGGAGACAGCATGAGAGTTCGCTGAGAGGTCTTAGAG 1260

Db 96 AACCCAGTCGCAAAATCAAAAGACATTCCTCCATCCCAAAACCCACCAAGCCCAAGCT 155
 Qy 41 ArgLeuProthlelleGlyHis***HisLeuLeuLysAspLysLeuLysHisValAla*** 60
 Db 156 CCGTCTCCCTTCATAGACACCTTCATCTCTTAAAGCAAACTCTCCATCGACATC 215
 Qy 61 IleAspLeuSerLysHisGlyProLeuPheSer*****PheGlySerMetProthr 80
 Db 216 ATCGACCTCTCCAAAACATAGTGCCTTATCTCTCTACTTGGCTCCAGCCAAAC 275
 Qy 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln****GluAlaThrSer 100
 Db 276 GTTGTGCTCTCCACACCAAAATTTCTCAAGCTCTTCTCCAAACGACAGAGCAATCTCC 335
 Qy 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla 120
 Db 336 TTCACACAGAGTTCCAAACCTCAGCAGCATAGACGCTCAGATGATGCTCAGTGGCC 395
 Qy 121 *****Pro***GlyProTyrTrp***PheValArgLysLeuLysMetAspLeuLeu 140
 Db 396 ATGTTCCCTTCGAGCTTACTGAGAGTTCTGAGAGAACTCTCATCATCAAGACCTTCTC 455
 Qy 141 AsnAlaThrThrValAsn***LeuArgProLeuArgThrGlnGlnLysLys***Leu 160
 Db 456 AACGCAACACGTAAACAGAGTGAAGCTTGAAGACCCAAACATCCGACATCTCTT 515
 Qy 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluLysLeuLeu 180
 Db 516 AGGTTATGCGCCCAAGCCCAAGAGCAAGAGCCCTTACTGATCCAGAGAGCTTCTG 575
 Qy 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGluLysLeuAsp 200
 Db 576 AATGACCAACAGACCATCTCATGATGATGCTCGGAGAGCTGAGAGATCATGAGAC 635
 Qy 201 IleAlaArgLysValLeuLysLysLys***GlyGluTyrSerLeuThrAspPheLys***Pro 220
 Db 636 ATGCTGCGAGAGTTCTTAAAGATCTTGGCAATACAGCTCATCACTTCACTCTGCGCA 695
 Qy 221 LeuLys***LeuLysValGlyLysTyrGluLysArgLysLeuAspLysLysLeuAsnLysPhe 240
 Db 696 TTGAAGCATCTCAAGGTTGGAAGATGAGAGAGATGACAGCATCTTGAACAAAGTTT 755
 Qy 241 AspProValValGluLysValLysLysLysArgArg***IleValArgArgArg***Asn 260
 Db 756 GACCTGTGCTTGAAGGCTCATCAAGAGCCGCTGAGATGCTGAGAGAGAGAGAGAAC 815
 Qy 261 GlyGlu*****GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
 Db 816 CGAGAGGTTGTGAGGTTGAGGTCACCGGGGTTTCTTGTACACTTCTTGAATTCGCT 875
 Qy 281 GluAspGluThr***GluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 300
 Db 876 GAGATGAGACATGAGATCAAAATCAACAGAGCAACATCAAGAGGCTTGTGTGAGC 935
 Qy 301 ***PheSerAlaGly***AspSerThrAla*****ThrGluThrAlaLeuAlaGluLeu 320
 Db 936 TTTTTCGACAGAAACGACTTCACAGCGGTGGCAACAGAGTGGGCAATGGGAGACTC 995

Qy 321 IleAsnAspPro***ValLeu*****AlaArgLysLys***TyrSerValAlaGlyLys 340
 Db 996 ATCAACAATCTTAAAGGTGTGTGAAGAGGCTGTGAGAGAGGTGTGAGAGAG 1055
 Qy 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360
 Db 1056 GACAGACTGTGAGAGATGACATCAAAACCTTCTTACATGAGCAATGAGAG 1115
 Qy 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys***GluLysCys 380
 Db 1116 GAGACATTCGCGCATGACCCGCCCTCCAGTGTGCAAAAGAAATGACAGAAAGTGT 1175
 Qy 381 ***IleAsnGly***Val***ProGluGlyAlaLeu*****PheAsnValIleProGlnVal 400
 Db 1176 GAGATTATGATATGATATCCAGAGAGGATGATTCCTTCAATGATATGAGCAAGTA 1235
 Qy 401 Gly***Asp***LysTyrTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu 420
 Db 1236 GGAGAGACCCCAAAATACGTGGAGACAGACATCGAGTTCGCTCTGAGAGGTCTTAGG 1295
 Qy 421 Thr***AlaGluGlyGluAla*****LeuAspLeuArgGly***HisPheGlnLeuLeu 440
 Db 1296 ACGAGGCTGAGAGGAGAGAGAGGAGGCTCTTGAATCTTGAAGAGCAAAATTTTCAACTTCTC 1355
 Qy 441 PropheGlySerGlyArg***MetCysProGlyVal***LeuAlaThrSerGly***Ala 460
 Db 1356 CCATTTGGGTCTGGAGAGAGATGTGCTTGAATGTGCTTACTGCTACTGCGAATGCGCA 1415
 Qy 461 ThrLeuLeuAlaSerLeuLysLysCysPheAspLeuGlnValLeuGlyProGlnGlyGln 480
 Db 1416 ACACTTCTGATCTTATTAATGATGCTTCTGAGAGTCTGAGAGTCTGAGAGAGAGCAG 1475
 Qy 481 IleLeuLysGly***AspAlaLysValSerMetGluLysArgAlaGlyLeuThrValPro 500
 Db 1476 ATATTGAAGGTGTGTACGCCAAAGTTAGCATGAGAGAGAGAGGCTTCACTGTTCGA 1535
 Qy 501 ArgAlaHisSerLeuValCysValProLeuAlaArgLysLysValAlaSerLysLeuLeu 520
 Db 1536 AGGCACTATGATCTTGTCTGTGTTCCACTTGCAGAGATCGAGGCTTGCATTAATCTCTT 1595
 Qy 521 Ser 521
 Db 1596 TCT 1598

RESULT 10
 AF022462 1824 bp mRNA linear PLN 02-MAR-1999
 LOCUS Glycine max cytochrome P450 monooxygenase CYP93C1 mRNA,
 DEFINITION complete cds.
 ACCESSION AF022462
 VERSION AF022462.1 GI:2739005
 KEYWORDS Glycine max.
 SOURCE Glycine max.
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Db 114 ACACCACTGCAAAATCAAAAGCATTTGGCATCTCCAAACCCACCAAGCCCAAGCCT 173

Qy 41 ArgLeuProPheIleGlyHis**HisLeuLeuYsaSprYsLeuLeuHisTyraIa*** 60
|||||
Db 174 CGCTTCCTTCATAGACACCTTCATCTTTAAAGACAATCTTCACATAGCAGCTC 233

61 IleAspIeuSerIySlySHicIyProIeuPheSer*****PheGlySerMetProThr 80

Db 234 ATGACCTCTCCAAAAACATGCTGCCCTATTCTCTCTACTTTGGCTCATGCCACC 293

Db 294 GTTGTGCTCACCACACGAAITGTTCAAGCTTCTCCTCCAAACGACGAGGCACTTC 352

```

Qy 101 Phe**ThrArgPheGlnThrSerAla**Arg**LeuThrTyrAsp*****ValAla 120
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 TTCAACACAGAGTTCAAAACCTCAGCATTAAAGACGCTCATGATGATAGCTCAGTGGCC 413

```

Dy 121 *****GlyProTyrTrp**PheValArgIysLeuIleMetAsnSpleuIeu 140
||| ||||||||| ||||||||| ||||||||| |||||||||

141 AsnAlaThrThrValAsn***LeuArgProLeuArgIleGlnGlnIleArgLys***Leu 166

[illegible]

Db 534 AGGGTTATGGCCCAAGCGCAGAGGACAGAGACCCCTTGCATTGACCGAGGACTTCTG 593

OY 181 LysTrp***AsnSerThr***SerMet***LeuGlyAlaClnIleArgasp 200
| | | | | | | | | | | | | | | | | | | | | |
Db 594 AATGGACCAAGACCATTCATCGATGTGATGCAGGAGCTGAGAGATCAGGAC 653

Qy 201 IleaIargGIuValleuIysile**GIyGIuTySerLeuThraSpheile**Pro 220

Db 654 ATGCGCTCGGAGGTTCTTAAGATCTTTGGCGAATACAGCCTCACTGACTTCACTGGCCA 713

Db 714 TTGACATCTCAGCITTGGAAGTAATGAGAGAGATCGACGACATCTTGAAACAATTC 773

Dy 241 AsproValValGluArgValIleLeuLysArgArg***IleValArgArgArg***Asn 266

Dd 774 GACCTGTGTTGAAGGGTATCAAGAAGCGCGTGAGTCGTGAGGAGAGAAAGAAC 833

261 GlyGlu*****GluGlyGlu***SerGlyVal**LeuAspThrLeuGluGluPheala 286

DB 834 GAGAGGICGTGAGGAGGICAGCGCGGCTTCCTCCTGACACCTTCCTGAAATTCCT 893

894 GAGATGAGACCATGAGATCAAAATCACCAAGACCAATCGAGGGCTTGTGTGCAC 95

Db 954 TTTTCTCGGAGGAACAGACTTCACACCGGTGGCAACAGAGTGGGCATTGGCAAGAACTC 101

QY 41 ArgLeuProPhe11eG1VH1s***H1sLeuLeuLysAspLysLeuLH1sTYRAla1*** 60
 DB 174 CTTCTTCCCTTCAATGACACCTTCACTTAAAGACAACTTCTCACTGACCTC 233
 QY 61 11eAspLeuSerLysLysH1eG1ProLeuPheSer****PheG1SerMetProThr 80
 DB 234 ATGACACTCTCCAAAACATGATGCTTATTTCTCTCACTTGGCTTCATCCAAAC 293
 QY 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln****GluAlaThrSer 100
 DB 294 GTTGTGCTTCCACACCAATTTCTCAAGCTTCTCCCAACGACAGAGCACTTCC 353
 QY 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyAsp*****ValAla 120
 DB 354 TTCACACACAGGTTCCAACTCAAGCCATPACACCTCACTATGATGCTCAAGTGGCC 413
 QY 121 *****Pro***GlyProTyTrp***PheValArgLysLeuL1MetAsnAspLeuLeu 140
 DB 414 ATGTTTCTTCCGACCTTCACTGAGAGTTCGAGAGAGCTCATCATGACAGACTTCCC 473
 QY 141 AsnAlaThrThrValAlaSn***LeuArgProLeuArgThrGlnGln11eArgLys***Leu 160
 DB 474 AAGGACACCACTGTAAACAAGTTAGGCTTTGAGAGACCAACACACCCGCAAGTTCTT 533
 QY 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluGluLeuLeu 180
 DB 534 AGGGATTATGCGCCACAGGCGACAGGACAGAGCCCTTGACTGACAGAGAGCTTCTG 593
 QY 181 LysTrp***AsnSerThr***SerMetMet***LeuGlyLysLysGluGlu11eArgAsp 200
 DB 594 AATATGACCAACAGCACATCTCATGATGATGCTCGGCGAGGCTGAGAGATCAAGAC 653
 QY 201 11eAlaArgGluValLeuLysL1e***GlyGluTySerLeuThrAspPheL1e***Pro 220
 DB 654 ATGCTTCCGACAGGTTCTTAAAGTTCTTGGCGAATACAGCTCACTGACTTCACTGGCCA 713
 QY 221 LeuLys***LeuLysValGlyLysTyArgLysArgG11eAspAsp11eLeuAsnLysPhe 240
 DB 714 TTGAAGCATCTCAAGGTTGAAAGTATGAGAGAGATGACAGACATCTTGAACAAAGTTC 773
 QY 241 AspProValValGluArgVal11eLysLysArgArg***11eValArgArgArg***Asn 260
 DB 774 GACCTGTCTGTAAGGCTCATCAAGAGCCGCTGATGATGAGAGAGAGAAAGAAC 833
 QY 261 GlyGlu*****GluGlyGlu***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
 DB 834 GAGAGGCTTGTGAGGAGGAGGTCACGCGGCTTCTTGTACACTTGTGTAATTCGT 893
 QY 281 GluAspGluThr***GluLysLysLysLysLysLysLysLysLysLysLysLysLys 300
 DB 894 GAGGATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACAT 953
 QY 301 ***PheSerAlaGly***AspSerThrAla*****ThrGluTrpAlaLeuAlaGluLeu 320
 DB 954 TTTTCTCGGACAGAAACAGATCTCAAGCGGTGCGAACAGAGTGGCATTTGCGCAAGCTC 1013
 QY 321 11eAsnAsnPro***ValLeu*****AlaArgGluGlu***TyrSerValGlyLys 340

DB 1014 ATCAACATCTTAAGTGTGTGAAAAGGCTCTGAGAGGCTCAAGCTTGTGTGGAAAG 1073
 QY 341 Asp***LeuValAspGluValAspThrGlnAsnLeuProTyTrp11eArgAla11eValLys 360
 DB 1074 GACAGCTTGTGAGAGAGTTGACATCAAAACCTTCTTACATTTAGAGCAATGTGAAG 1133
 QY 361 GluThrPheArgMetH1sProLeuProValValLysArgLysCys***GluGluCys 380
 DB 1134 GAGACATTCGCGATGACACCCGCTCCAGTGTGCAAAAGAAAGTGCACAGAAAGTGT 1193
 QY 381 ***11eAsnGly***Val***ProGluGlyAlaLeu*****PheAsnValTrpGlnVal 400
 DB 1194 GAGATTAAATGATATGTATCCAGAGGAGGATTTGATCTTCAATGATAGCAAGTA 1253
 QY 401 Gly***Asp***LysTyTrpAspArgProSerGlu***ArgProGluArgPheLeuGlu 420
 DB 1254 GGAAGAGACCCCAATATCTGAGACAGACATGAGATTCCTGAGAGGTTCTTAGAG 1313
 QY 421 Thr***AlaGluGlyGluAla*****LeuAspLeuArgGlyLys***H1sPheGlnLeuLeu 440
 DB 1314 ACAGGAGCTGAGAGGAGAGAGGAGGCTTGTATCTTATGAGGAGCAATTTCAACTTCTC 1373
 QY 441 ProPheGlySerGlyArg***MetCysProGlyVal***LeuAlaThrSerGly***Ala 460
 DB 1374 CCAATTTGGTCTGAGAGAGATGTGCTCCATGATGTGCTTACTTCCGAAATGCGCA 1433
 QY 461 ThrLeuLeuAlaSerLeuL1eGlnCysPheAspLeuGlnValLeuGlyProGlnGlyGln 480
 DB 1434 ACACCTTGTGATCTTATTAATCACTGCTTCACTTCAAGTGTGCTGCAAGAGCAG 1493
 QY 481 11eLeuLysGly***AspAlaLysValSerMetGluGluArgGlyLysLeuThrValPro 500
 DB 1494 ATTTGAAGGCTGTGACCCCAAGATTAGCATGAGAGAGAGAGCCGCTCACTGTTCCA 1553
 QY 501 ArgAlaH1sSerLeuValCysValProLeuAlaArgG11eGlyValAlaSerLysLeuLeu 520
 DB 1554 AGGAGCATATGCTGTCTGTGTTCACCTTGCACAGCATCGGCTTGCATTTAACTCTT 1613
 QY 521 Ser 521
 DB 1614 TCT 1616

RESULT 12
 AF195819
 LOCUS AF195819 1902 bp DNA linear PLN 23-MAR-2000
 DEFINITION Glycine max isoflavone synthase 2 (ifs2) gene, complete cds.
 ACCESSION AF195819
 VERSION AF195819.1 GI:7288454
 KEYWORDS
 SOURCE Glycine max.
 ORGANISM Glycine max.
 Eukaryota; Vitisidiantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE
 1 (bases 1 to 1902)

AUTHORS Jung, W., Yu, O., Lau, S.M., O'Keefe, D.P., Odell, J., Fader, G. and McConigle, B.
TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes
JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)
MEDLINE 20124255
PUBMED 10657130
REFERENCE 2 (bases 1 to 1902)
AUTHORS Jung, W., Yu, O., Odell, J., Fader, G. and McConigle, B.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-1999) Nutrition and Health, Dupont, PO Box 80402, Wilmington, DE 19880-0402, USA
FEATURES
Source Location/Qualifiers
1..1902
/organism="Glycine max"
/db_xref="taxon:3847"
1..1902
/gene="ifg2"
join(1..948,1085..1902)
/gene="ifg2"
/product="isoflavone synthase 2"
join(52..948,1085..1753)
/gene="ifg2"
/note="cytochrome P450"
/codon_start=1
/product="isoflavone synthase 2"
/protein_id="AAF4514.1"
/db_xref="GI:7288455"
/translation="MLELALGLVLAFLHRTPTAKSARLHPSPSPKRLPF
IGHHLKDKLHYALIDLSKKQPLFSLYQSMPTVAATPELFLQTHATSFN
TRFQSAIRLTYDSSVAMPFGPYKPVKRLIMNDLPNATVNLRLPQTRKFL
RYMAQGAQKPLDLTEELIKMTNSTISMMILGEARELDIAEVLKIFGYSILDFI
WPLKHLAVGKYEKRIDIDLNKPDVVERVIRKRIYRRKNGVEGEGSVGLDTL
LEFADETMEIKITKDHIEGLVDFPSAGTSTAVATEMALMLNNKYLAKAREEV
YSVGGKRLVDEVDTQNLPIYKALVKEITRMHPLPVKKKCTECCEILINGVILPQAL
ILFNWQVGRDPKWDPRSEPRFLFTGAGGEGPLDRQGHQQLPFGSGRMCPL
GVNLATSGMATLASLIQCFLDVLGPOGQILKGGDAKYSMEBRAGLTVPRASHIVCV
PLARIGVASKILLS"

BASE COUNT 517 a 455 c 440 g 490 t
ORIGIN

Alignment Scores:
Pred. No.: 1.06e-251 Length: 1902
Score: 2201.50 Matches: 451
Percent Similarity: 79.72% Conservative: 1
Best Local Similarity: 79.54% Mismatches: 69
Query Match: 91.88% Indels: 46
DB: 8 Gaps: 1

US-09-857-581-66 (1-521) x AF195819 (1-1902)

QY 1 MetLeuLeuGluLeuAlaLeuGlyLeu***ValLeuAlaLeuPhe***HisLeuArgPro 20
DB 52 ATGTTGCTGACTGACCTGTTATGTTGGCTGCTGTTCTGACTTGCACGCGCC 111
QY 21 ThrPro***Ala***SerIysAlaLeuArgHisLeuProAsnProProSerPro***Pro 40
||||| ||| ||||||||||||||||||||||||||||||||||| |||

DB 112 ACNCCCATGCAAAATCAAAAGCACTTCGCATCTCCCAAAACCAAGCCCAAGCT 171
QY 41 ArgLeuProPheIleGlyHis***HisLeuLeuAspIysLeuLeuHisTrpAla*** 60
DB 172 CGCTTCCTCCCTCAATAGCAACCTTCATCTCTTAAAGCAAACTTCCTCAGCTGCACTC 231
QY 61 IleAspLeuSerIysIysHisGlyProLeuPheSer*****PheGlySerMetProThr 80
DB 232 ATGACCTCTTCCAAAMACATGCTCCCTTATTCCTCTTACTTGGCTCCATCCCAACC 291
QY 81 ValValAlaSerThrProGluLeuPheIysLeuPheLeuGln*****GluAlaThrSer 100
DB 292 GTTGTGCTGCTCCACACCAATGTTCAAGCTCTCTTCCAAAGCAAGGCAAGCTTCC 351
QY 101 Phe***ThrArgPheGlnThrSerAla***Arg***LeuThrTyrAsp*****ValAla 120
DB 352 TTCACACAGAGTTCCAAACCTCAGCATTAAGACCTCATATAGCTCATGAGGACC 411
QY 121 *****Pro***GlyProTyrTrp***PheValArgIysLeuIleMetAsnAspLeuLeu 140
DB 412 ATGCTTCCTCCGACCTTACCTGAGATTCGTGAGAGAGCTCATGACACACTTCCC 471
QY 141 AsnAlaThrThrValAsn***LeuArgProLeuArgThrGlnGlnIleArgIys***Leu 160
DB 472 AACGCCACACACTTAAACAACTTGAAGGCTTGGAGACCCCAAGACCCGCAACTTCTT 531
QY 161 Arg***MetAlaGln***AlaGluAla***LysProLeuAsp***ThrGluLeuLeu 180
DB 532 AGGCTTATGCGCCCAAGCGCAGAGGCAAGAGCCCTTACTGATCCAGAGACTTCTG 591
QY 181 IysTrp***AsnSerThr***SerMetMet***LeuGlyGluAlaGluIleArgAsp 200
DB 592 AATGGACCAACAGCAACATCTCATATGATGCTCGCGAGGCTGAGAGATCAAGAGAC 651
QY 201 IleAlaArgGluValLeuLysIle***GlyGluTyrSerIleThrAspPheIle***Pro 220
DB 652 ATCGCTCCCGAGGTTCTTAAGATCTTGGCGAATCAACCTCATGACTTCACTGCGCA 711
QY 221 LeuLys***LeuLysValGlyLysTyrGluLysArgGlnIleAsnAspIleLeuAsnLysPhe 240
DB 712 TTGAAGCATCTCAAGGTTGAAAGTATGAGAGAGATGACGACATCTTCAACAGCTTC 771
QY 241 AspProValValGluArgValIleLysLysArgArg***IleValArgArgArg***Asn 260
DB 772 GACCTGCTGCTTGAAGAGGTCATCAAGAGCGCGTGAATCTGAGAGAGAGAGAAC 831
QY 261 GlyGlu*****GluGlyGluLys***SerGlyVal***LeuAspThrLeuLeuGluPheAla 280
DB 832 GAGAGAGTTGTAGGGTGAAGGTGAGGCGGGGTTTCTTGACACTTGGCTGAATTCGCT 891
QY 281 GluAspGluThr***GluIleLysIleThrLys*****IleLysGlyLeuValVal--- 299
DB 892 GAGAGATGAGACATGAGATCAAAATCAACCAAGAGCAACATGAGGCTCTTGTTGCTG 951
QY 299 ----- 299
DB 952 AGTTCTGCTCATTCATTCATGCAAAATATGCAAGATTTTGTAAACAAGATCGAGAA 1011

QY 299 ----- 299

Db 1012 TTGACATTATATATCATGTGGTGCATTAAATTAACGGTACGATCTTAAATTCATAT

QY 300 ----- Asp***PheSerIaGly***AspSerThrIaA*****ThrGluTr 315

Db 1072 TGTGATGTGACAGAGCTTTTCTCGGACAGAACAGACTCCACAGCGGTGGACACAGATG 1111

QY 315 pAlaLeuAaGluLeuIleAaAaenPro***ValIeu*****AaArgGluGlu***Ty 335

Db 1132 GCGATTGGACAGAACTCATCAACATCTTAAGGTGTGAAAGGCGTCTGAGAGGTCTA 1191

QY 335 rSerValaIaGlyIyAsp***LeuValaSpGluValaAspThrGlnAaenLeuProTyr11 355

Db 1192 CAGTGTGTGGGAAAGACAGACTGTGACAGAGTTGACATCAAACTTCTTACAT 1251

QY 355 eArgIaIaIeValIyGluThrPheArgMetHAProProLeuProValIaIyArgLy 375

Db 1252 TAGACCAATGCTGACAGAGACATTCGCAATGCAACCCGCACTCCAGTGTCAAAAGAA 1311

QY 375 sCys***GluGluCys***IleAaenGly***Val***ProGluGlyIaIeAaen***Ph 395

Db 1312 GTGCACAGAAAGTGTGAGATTATGATATGATATCCAGAGAGCATTTGATCTCTT 1371

QY 395 eAaenValITrGlnIaGly***Asp***LysTyrITrAspArgProSerGlu***ArgPr 415

Db 1372 CATGTATGGCAAGTGTGAGAGAGACCCCAATATCTGGGACAGACATCGAGTTCCGTCC 1431

QY 415 oGluArgPheLeuGluThr***AlaGluGlyGluAa*****LeuAaLeuArgGly*** 435

Db 1432 TTAGAGGTTCTCTAGAGACAGCGGCTGTAGAGGAGAGCGGCTCTTGTATCTTAGGGGACA 1491

QY 435 *HisPheGluLeuLeuProPheGlySerGlyArg***MetCysProGlyVal***IeAa 455

Db 1492 ACATTTTCAACTTCTCCATTTGGGCTGTGGAGAGATGTGCTCTGAGTCAATCTGAC 1551

QY 455 aThrSerGly***AlaThrLeuLeuAaSerLeuIleGlnCysPheAaLeuGlnValIe 475

Db 1552 TACTTCGGGAAAGGACACATCTTGTGATCTTATTCAGTGTCTTCACTTGCAGAGTCTT 1611

QY 475 uGlyProGlnGlyGlnIleLeuLysGly***AspAlaIyValaSerMetGluGluArgA1 495

Db 1612 GCGTCAACAGACAGATATTGAAGGAGTGTGACGCCAAAGTTAGCATGAGAGAGAGC 1671

QY 495 aGlyLeuThrValProArgAlaHisSerLeuValCysValProLeuAaAaArgIleGlyA 515

Db 1672 CGGCTCACTGTCCAGAGGACATAGTCTGTGTGTGATTCACATTCGAGATCGGCGT 1731

QY 515 IAlaSerLysLeuLeuSer 521

Db 1732 TGCATCTAAACTCTTCTT 1750

RESULT 13

LOCUS AF195818 1800 bp DNA linear PLN 23-MAR-2000

DEFINITION Glycine max isoflavone synthase 1 (ifsl) gene, partial cds.

ACCESSION AF195818

VERSION AF195818.1 GI:7288452

KEYWORDS

SOURCE Glycine max.

ORGANISM Glycine max

REFERENCE 1 (bases 1 to 1800)

AUTHORS Jung, W., Yu, O., Lau, S.M., O'Keefe, D.P., Odell, J., Fader, G. and McConigle, B.

TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes

JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)

MEDLINE 20124255

PUBMED 10657130

REFERENCE 2 (bases 1 to 1800)

AUTHORS Jung, W., Yu, O., Odell, J., Fader, G. and McConigle, B.

TITLE Direct Submission

JOURNAL Submitted (18-OCT-1999) Nutrition and Health, Dupont, PO Box 80402, Wilmington, DE 19880-0402, USA

FEATURES

source

1..1800

location/Qualifiers

gene

/organism="Glycine max"

/db_xref="taxon:3847"

1..>1800

/gene="ifsl"

join<1..893,1112..>1800

/gene="ifsl"

/product="isoflavone synthase 1"

join<1..893,1112..>1800

/gene="ifsl"

/note="cytochrome P450"

/codon_start=3

/product="isoflavone synthase 1"

/protein_id="AAF45142.1"

/db_xref="GI:7288453"

/translation="LEIAGLFTVALFLHPTSAKSKALHLNPPSPKPLPTIG
HHHLKRLHVALIDLSKKHGLFSLSGSPFTVAYSTELFKFLDTHEATSPNTR
FQTSAIRRLTYDNSVAMVFGPYKFKVRLIMNDLNAVTVNKLPLTQIRKFLRV
MQSAEIQKLDVTEBELKMTNSTISMMLGASEIRDLARVLKIPGYSGLDPTLP
LKYLVKYEKRLIDILINKFDPVVERIVIKRREIVRRKNGEVEGASGVPLDTLLE
FADETWELITKEQIKGLVVDPSAGDSTAVATVETAALELNNRVLQKAREEYS
VGDGRLVDEVDTQNLPIYRAIYKTFPRHPLPLVVRKCTECCINQYVPEGALVL
FVWVGAGDVKYWRSPSEFRPERPLETGAEGAGLDLRQGHPLLPFGSGRMKCGV
NATSGAATLIALSLIQCFDLQVLPQGIILKGDADAKVMEERAGLTVRAHSLVCVPL
ARIGVASKSLLS"

BASE COUNT 476 a 441 c 413 g 470 t

ORIGIN

Alignment Scores:

Pred. No.: 3.04e-250 Length: 1800

Score: 2189.00 Matches: 451

Percent Similarity: 76.18% Conservative: 0

Best Local Similarity: 76.18% Mismatches: 68

Query Match: 91.36% Indels: 74

DB:	8	Gaps:	1
US-09-857-581-66 (1-521) x AF195818 (1-1800)			
Qy	3	LeuGIuLeuAlaLeuGIyleu***ValLeuAlaLeuPhe***HisLeuArgProThrPro	22
Db	3	CGGAACTTCGACCTTGTTGTTGTTAGCTTTGTTTCGACTTCGCTCCCAACCA	62
Qy	23	***Ala***SerIyAlaLeuArgHisLeuProAsnProSerPro**ProArgLeu	42
Db	63	ACTGCAAAATCAAAAGACCTTCGCCACCTCCCAACCCCTCAAGCCCAAGCCCTGCTT	122
Qy	43	ProheliLeuIyHis***HisLeuLeuIyAspIySleuLeuHisIyTrAla***IleAsp	62
Db	123	CCCTTCATTTGGCCACCTTCACCTCTTAAAGATAACTTCCTCACTAGCACTCACTGAT	182
Qy	63	LeuSerIyIyHisGIyProLeuPheSer*****PheGIySerMetProThrVal	82
Db	183	CTCTCCAAAAGATGGCCCTTATTCCTCTCTCTCTGAGCTCATCCCAACCGTCTT	242
Qy	83	AlaSerThrProGIuLeuPheIySleuPheLeuGIu****GIuAlaIthrSerPhe**	102
Db	243	GGCTCAACCCCTGATGTTTCAAGTCTCTTCCTCAACCCAGAGGCAACTCTTCAAC	302
Qy	103	ThrArgPheGIuThrSerAla***Arg***LeuThrTyAsp*****ValAla*****	122
Db	303	ACAGAGTTCAAAACCTCTGCATATAGACCCCTCACTTACGACACTCTGTGGCATGTT	362
Qy	123	Pro***GIyProTyTrTP**PheValArgIySleuIleMetAsnAspLeuAsnAla	142
Db	363	CCATTTCGACCTTACGAGAGCTTCGAGAGAGCTCATCATACAGACTTTCACACGC	422
Qy	143	ThrThrValAsn**LeuArgProLeuArgThrGIuGIuIleArgIy***LeuArg**	162
Db	423	ACCAACCTCAACAAGCTCAAGGCTTTGAGAGACCCAAAGATCCGAGTTCTTAGGTT	482
Qy	163	MetAlaGIu***AlaGIuAla***IySProLeuAsp**ThrGIuGIuLeuLeuIySTP	182
Db	483	ATGGCCCAAGCGCAGAGGCCCAAGAGCCCTTGACGTCAAGAGAGACTTTCAAATGG	542
Qy	183	***AsnSerThr***SerMetMet***LeuGIyGIuAlaGIuGIuIleArgAspIleAla	202
Db	543	ACCAACAGCAACATCTCCATGATGATGCTCGGCGAGCGTAGAGATCAAGACATCGCT	602
Qy	203	ArgGIuAlaLeuIyIle***GIyGIuIySerIleuThrAspPheIle***ProIeuIyS	222
Db	603	GGCGAGGTTCTTAAATCTTGGCGAATACAGCTCACTGACTTCATTCGCTTTGAAG	662
Qy	223	**LeuIySValGIyIyGIyTrGIuIySArgIleAspAspIleLeuAsnIySPhAspPro	242
Db	663	TATCTCAAGGTTGAAAGTATGAGAAAGAGATGATGACATCTTGAAACAGTTCCACCT	722
Qy	243	ValValGIuArgValIleIySlySArgArg**IleValArgArgArg**AsnGIyGIu	262
Db	723	GTGCTGAAGGGGTGATCAAGAGAGCCGCTGAGATCGTCAGAAAGAAAGAAAGAGAA	782
Qy	263	****GIuGIyGIuIy***SerGIyVal***LeuAspThrLeuLeuGIuIlePheAlaGIuAsp	282
Db	783	GTTCATGAGGCGAGGCGAGGCGGTCTTCTCGACACTTTCGTAATTCGCTGAGAC	842
Qy	283	GIuThr***GIuIleIySleIthrIyS****IleIySlyLeuValVal-----	299
Db	843	GAGACATGAGATCAAAATATACCAAGAGCAAAATCAAGGGCTTGTGT-CGTAAGTTT	901
Qy	299	-----	299
Db	902	CCCTTCCTCTCTACTTATTTACTTTCCTTTCATTCATTCATATGATATGCGATTAAATAG	961
Qy	299	-----	299
Db	962	TATACTATATGAAAAATATATATACGACTACAGGTGAAGATATGTGTGTTTTTTA	1021
Qy	299	-----	299
Db	1022	AAAAGATACAGAAAGTTCCTTTATATGATATGATATAGATATATTCATCAAGTGAA	1081
Qy	300	-----Asp***PheSerAlaGIy***AspSerThr	309
Db	1082	ACTATATATTCCTATTTTGATATGATAGACTTTCCTGACAGGACAGATCCACA	1141
Qy	310	Ala*****ThrGIuTrPAlaLeuAlaGIuLeuIleAsnAsnPro***ValLeu*****	329
Db	1142	CGCGTGCAACAGAGTGCGGCAATGGGAGGCTCATCAACATCCAGGGGTGTGCAAAAG	1201
Qy	330	AlaArgGIuGIuIy***TySerValValGIyIySAsp***LeuValAspGIuValAspThr	349
Db	1202	GCTCGTGAGAGGTCTACAGATGTTGTGGCAAAAGATAGACTGTGGAGAGGTTCACCT	1261
Qy	350	GIuAsnLeuProTyTrIleArgAlaIleValIySGIuIyThrPheArgMetHisProProLeu	369
Db	1262	CAAAACCTCTTCACTTATGAGGCGATTTGAAAGAGACATTCGATGACACCCACATC	1321
Qy	370	ProValIyIySArgIyCys***GIuIyCys***IleAsnGIy***Val***ProGIu	389
Db	1322	CCAGGTGTCAAAAGAAATGACACAGAGAGTGTGATATATGCTATGTGATCCAGAG	1381
Qy	390	GIyAlaLeu*****PheAsnValTrPGIuValGIy***Asp***TyTrTyTrPAspArg	409
Db	1382	GGAGCATTTGTTCTTTTCAAGTTTGCAAGTATGAGAGGAGACCCCAATATCTGGACAGA	1441
Qy	410	ProSerGIu***ArgProGIuArgPheLeuGIuIyThr***AlaGIyGIuAla*****	429
Db	1442	CCATAGAAATTCCTGCCGAAGAGTCTTTAGAAATGCTGTCTGAAGGGAAGCAGGGCCT	1501
Qy	430	LeuAspLeuArgGIy***HisPheGIuLeuLeuProPheGIySerGIyArg***MetCys	449
Db	1502	CTTATGCTTGAAGGCGACCATTTCCAACTCTCCCATTTGGCTGGGAGAGATATGTC	1561
Qy	450	ProGIyVal***LeuAlaIthrSerGIy***AlaIthrLeuLeuAlaSerLeuIleGIuIyS	469
Db	1562	CCGTGTGCAATTTGGCTTACTTCAGAAATGGCAACACTTCTTGCAATCTTATTCATATC	1621
Qy	470	PheAspLeuGIuIyLeuGIyProGIuGIuIyIleLeuIyGIy***AspAlaIySVal	489

QY 15 Phe***HisLeuArgProThrPro***Ala***SerValAlaLeuArgHisLeuProAsn 34
 Db 3 TTCTGCACCTGCTGCCACCAAGTGCMAAATCAAGACCTTCGCCACCTCCCAAC 62
 QY 35 ProProSerPro***ProArgLeuProPheIleGlyHis***HisLeuLeuValAspLys 54
 Db 63 CCCCCAGGCCCAAGCTGCTCTCCCTTCATGCGCACTTCCTTAAAGATTA 122
 QY 55 LeuLeuHisTyrAla***IleAspLeuSerLysHisGlyProLeuPheSer***** 74
 Db 123 CTTCACCATATGACATCATGATCTCTCAAAAAGCATGGCCCTTATCTCTCTCC 182
 QY 75 PheGlySerMetProThrValValAlaSerThrProGluLeuPheLysLeuPheLeuGln 94
 Db 183 TTCGGCTCCATGCGCAACCGGTGCTGCCACCCCTGATGTTGTCACCTTCCTCCAA 242
 QY 95 *****GluAlaThrSerPhe***ThrArgPheGlnThrSerAla***Arg***LeuThr 114
 Db 243 ACCCAAGAGCAACTTCCTTCACCAAGTTCACCACTTCGCCCAAGACCTCCACT 302
 QY 115 TyrAsp*****ValAla*****Pro***GlyProTyrTyr***PheValArgLysLeu 134
 Db 303 TACCAACAATCTGTGGCCATGCTTCATTCGACCTTACGAGGTTGTGAGGAAGCTC 362
 QY 135 IleMetAsnAspLeuLeuAsnAlaThrThrValAsn***LeuArgProLeuArgThrGln 154
 Db 363 ATCATGAAGACCTTCACACGCCACCAACCGTCAACAGCTGACGGCTTTGAGAACCCA 422
 QY 155 GlnIleArgLys***LeuArg***MetAlaGln***AlaGluAla***LysProLeuAsp 174
 Db 423 CAGATCCGCAAGTCTTAAAGGTTATGGCCCAAGCCCAAGGCCCAAGCCCTTGAC 482
 QY 175 ***ThrGluGluLeuLeuLysTyr***AsnSerThr***SerMetMet***LeuGlyGlu 194
 Db 483 GTCACGAGAGAGCTTCTCAATATGACCAACACACACATCTCCATGATGATGCTGCCAG 542
 QY 195 AlaGluGluIleArgAspIleAlaArgGluValLeuLysIle***GlyLysTyrSerLeu 214
 Db 543 GCTAGAGATGACAGACATCGCTCCGAGGTTCTTAAGATCTTGGGCAATACAGCTTC 602
 QY 215 ThrAspPheIle***ProLeuLys***LeuLysValGlyLysTyrGlyLysArgIleAsp 234
 Db 603 ACTGACTTATCTGGCCTTGAAGTATCTCAAGGTTGGAAATATGAGGAAGAGATGAT 662
 QY 235 AspIleLeuAsnLysPheAspProValValGluArgValIleLysLysArgArg***Ile 254
 Db 663 GACATCTTGAACAGATTCGACCTGCTGAAAGGTCATCAAGAGCCCGTGGAGATC 722
 QY 255 ValArgArgArg***AsnGlyGlu*****GluGlyGlu***SerGlyVal***LeuAsp 274
 Db 723 GTCAGAAAG 782
 QY 275 ThrLeuLeuGluPheAlaGluAspGluThr***GluIleLysIleMetLys*****Ile 294
 Db 783 ACTTCTGATTCGCTGAG 842

QY 295 LysGlyLeuValValAlaAsp***PheSerAlaGly***AspSerThrAla*****ThrGlu 314
 Db 843 AAGGACCTTGTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 902
 QY 315 TrpAlaLeuAlaGluLeuIleAsnAsnPro***ValLeu*****AlaArgGluLys*** 334
 Db 903 TGGCATTTGCGACAGCTTCATCAACATCCAGGAGGTTGCAAAAGGCTGGAGAGGCTC 962
 QY 335 TyrSerValValGlyLysAsp***LeuValAspGluValAlaAspThrGlnAsnLeuProTyr 354
 Db 963 TACAGTGTGGGCAAAAGATATGACTGTTGACAGAGTATGACATCAAAACCTTCCTTC 1022
 QY 355 IleArgAlaIleValLysGluThrPheArgMetHisProProLeuProValValLysArg 374
 Db 1023 ATTAAGGCGCATTTGAGAGAGACATTCGATGACACCACTCCGATGGTCAAAAGA 1082
 QY 375 LysCys***GluGluLys***IleAsnGly***Val***ProGluGlyAlaLeu***** 394
 Db 1083 AAGTCAACAGAGAGAGTGAAGATTAAATGGATTAATGATTCACAGAGAGAGATTCGTTCTT 1142
 QY 395 PheAsnValTrpGluValAlaGly***Asp***LysTyrThrAspArgProSerGlu***Arg 414
 Db 1143 TTCAATGTTGGCAAGTGAAG 1202
 QY 415 ProGluArgPheLeuGluLysThr***AlaGluGlyGluAla*****LeuAspLeuArgGly 434
 Db 1203 CCGAGAGAGTCTTAAGAACTGTCGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1262
 QY 435 ***HisPheGlnLeuLeuProPheGlySerGlyArg***MetCysProGlyVal***Leu 454
 Db 1263 CAGCATTTCCAACTCTCCCACTTGGGCTGGAGAGAGAGATGGCCCTGCTGATTCATTTG 1322
 QY 455 AlaThrSerGly***AlaThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnVal 474
 Db 1323 GCTACTTCAGAAATGGCAACACTTTCGATCTCTTATCCAAATGCTTTGACCTCGCAAG 1382
 QY 475 LeuGlyProGlnGlyGlnIleLeuLysGly***AspAlaLysValSerMetGluGluArg 494
 Db 1383 CTGGGCTTCAGAGACAAATTTCAAAAGGTATGATGATGACAAAGTTGATGAGAGAGCA 1442
 QY 495 AlaGlyLeuThrValProArgAlaHisSerLeuValCysValProLeuAlaArgIle 513
 Db 1443 GCTGGCTTCAGAGTTCAGAGGAGCAATATGCTGTTGTGTTTCACCTTCAGAGATC 1499

Search completed: May 18, 2003, 13:07:12
 Job time : 2876 secs

GenCore version 5.1.4_p5_4578
 Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
 Run on: May 18, 2003, 07:48:03 : Search time 239 Seconds
 (without alignments)